Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 27 (2024), No. 8, pp. 2301–2336
DOI : 10.47974/JDMSC-1859

Understanding user polarisation regarding COVID-19 vaccines through social network analysis

Harshit Garg *
Srishti Jhunthra †
Department of Computer Science & Engineering
Bharati Vidyapeeth's College of Engineering
New Delhi 110063
India

Rahul Goel §
Shakshi Sharma ‡
Institute of Computer Science
University of Tartu
Narva mnt 18
Tartu 51009
Estonia

Vedika Gupta @
Jindal Global Business School
O. P. Jindal Global University
Sonipat 131001
Haryana
India

Rajesh Sharma #
Institute of Computer Science
University of Tartu
Narva mnt 18
Tartu 51009
Estonia

* *E-mail:* gargharshit2000@gmail.com (Corresponding Author)
† *E-mail:* jhunthrasrishti@gmail.com
§ *E-mail:* rahul.goel@ut.ee
‡ *E-mail:* shakshi.sharma@ut.ee
@ *E-mail:* vgupta2@jgu.edu.in
# *E-mail:* rajesh.sharma@ut.ee

Pranav Dass [$]
Department of Computer Science
Shyam Lal College
University of Delhi
Delhi 110032
India

Abstract

Whenever any critical incident comes into the limelight, people tend to discuss it over social media platforms to exchange ideas, perspectives and opinions globally. Sometimes the chain of discussions lead to generation of new perspectives and enhance the general understanding of short-term or long-term implications of the problem at hand. Social media platforms have been an important tool to study and analyse opinions of stakeholders, consumers, citizens, customers, or clients. If positive, these opinions can help gain confidence for a business product or policy law. If otherwise the negative opinions can help the business, or the government understand online social media users about their feelings and take the necessary steps to counter bad reputation and promote customer satisfaction. Having an in-depth descriptive understanding of how the state of things are, leads to a realization of what circumstantial strategies should be adopted or followed to tackle such scenarios. This paper presents the as-is scenario of discussions held on social media platform Twitter regarding the various vaccines for COVID-19, which has been a critical topic of interest worldwide.

This study proposes a detailed analysis of the variety of perspectives and user polarisation persisting on Twitter in different dimensions. We segment Twitter users into two categories proponents (those propelling the idea of vaccination) and opponents (those opposing the idea of vaccination). Further, we have also categorized the Twitterati (Twitter users) on the basis of information seekers (inquisitive), information providers (informative) and opinion providers (opinionated). The analysis shows us the change of thoughts and opinions of the twitter users over a period of time. Overall, the paper presents quantitative and qualitative analysis depicting some analytical views taking different parameters into consideration.

Subject Classification: *05C82 (Graph Theory), 92B20 (Neural Networks), 68U15 (Computing methodologies for text processing).*

Keywords: *Twitter, COVID-19, Vaccinations, Natural language processing, Networks, Social media.*

1. Introduction

Beyond impacting individual health and physical well-being, the COVID-19 pandemic, which began in December 2019[1], has had a negative

[1] https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[$] *E-mail:* **pdasscs@shyamlal.du.ac.in**

impact on society as a whole in various ways [40, 41, 42]. The continuous discussion on social media has reinforced the worries, apprehensions, and doubts around COVID-19 and the currently available vaccinations, resulting in an infodemic. The virus, which was previously unknown and often developing and unsettling social and political responses to it, has created a highly unpredictable atmosphere that lead to the eruption of massive influx of opinions on social media. The social media witnessed a divide of users when it came to their viewpoints on different topics related to COVID-19 and its vaccines [36].

Given the constant flood of opinions and the massive influence on society as a whole, it became intriguing to study the user polarisation on social media caused by the COVID-19 outbreak in an automated fashion [38]. Along this line, we propose a labelled COVID-19 vaccine-related dataset, along with analyses to shed light on the discussions and propagation of user opinions.

The earlier studies in this regard, put a broad emphasis by focusing on a single technique. One type of analysis is, network-based approach where researchers have performed network visualization techniques to study the trend or formation of echo chambers on social media [6, 15, 20]. The analysis [6] discovered echo chambers in the Italian vaccine debate on Twitter. The network polarization technique [20] was utilised to study the trend and examine each new vaccination that was developed. Another sort of analysis is the use of natural language processing-based approach where the researchers in COVID-19 debates to prevent misinformation. Sentiment analysis [11, 19, 32] and topic analysis [12, 13, 33] are two NLP approaches that have been frequently used to identify current discussions and user sentiments.

Specifically, the analysis of text-based elements such as sentiments retrieved from texts, together with the identification of the communities created by various sorts of users, reveals interesting patterns of such users [34]. Furthermore, in the year 2020, when the most recent vaccinations were released by various research institutes throughout the world, individuals began tweeting their thoughts and ideas on various vaccines [2,32]. In this paper, we collected and analysed tweets related to five different vaccines- Covaxin, Covishield, Moderna, Pfizer and Sputnik. Overall, more than 204,000 tweets have been collected, spanning across three months- February, March and April 2021. Our main focus is to examine opinions on various vaccination kinds that can provide us with a better understanding of how vaccines have influenced user opinions online. We use a transfer learning approach to classify our gathered

COVID-19 vaccinations dataset. Then, we employ a variety of analysis (including network and NLP-based methods) to figure out what patterns different sorts of people use. Finally, we wrap up our research and discuss the work's future directions and constraints.

The rest of the paper is organised as follows. Section 2 presents the related work. Section 3 describes the dataset. Section 4 discusses around dataset curation. Section 5 presents the networks thus formed. Section 6 describes the results for engagement analysis and emotion analysis. Section 7 draws the conclusion.

2. Related Works

This section discusses related work from two perspectives: network-based approaches and natural language processing-based approaches.

2.1 *Approaches Based on Networks*

In literature, a lot of researchers have performed various network visualization techniques to study the trend or formation of echo chambers on social media [6, 15, 20]. [6] proposed a study of the formation of an echo chamber on the debate that occurred in the Italian vaccination, which took place on Twitter. The study showed the formation of networks formed among different users based on their retweets, mention, and followers. [31] Identified two or more communities on another famous social media platform, Reddit. The study showed the formation of various networks among different users worldwide. They proposed collaborative, social computing, and many more techniques to analyze multiple factors that can play a vital role in the identification of communities between Reddit users. The study thereby concludes that social media has a key role in opinion sharing. Thus, the formation of echo chambers net- works can be possible by studying and analyzing various situations upon which the user reacts. Many researchers used the network polarisation technique [27] to identify users' communities. [20] used this technique to study the trend and analyze every new vaccine that comes into existence. Whereas, in the past [15] also used a similar network visualization technique to generate connections between various sentiments and emotions of different users with a political aspect into consideration. Thus, it can be concluded that communities, with the help of network visualization, can be seen as a method to find the existence of a relationship between different categories, which would help in the identification of the ongoing trends [35, 39].

2.2 *Approaches Based on Natural Language Processing (NLP)*

Previous research studies have also presented Natural Language Processing's capabilities and limits for detecting misinterpretation of information, also known as misinformation[25] or abusive language detection [44]. It lays up a theoretical foundation for comprehending various forms of misinformation- rumours, fake news, and other ongoing issues, for example. Researchers in COVID-19 debates have frequently used NLP in order to prevent misinformation [1, 11, 19, 30].

[30] employed a natural language processing (NLP) technique to determine the top eleven primary themes of conversation about COVID-19 and the sentiments of Twitter users in the COVID–19 pandemics. The study explains the function and relevance of social media, which has a significant impact on users. Similarly, [19] presented how global sentiments about the pandemic crisis are reflected on Twitter. Specifically, in the Covid - 19 pandemic, the study focused on four primary emotions: fear, anger, sadness, and joy. The study suggests that the public is experiencing certain collective concerns as a result of the COVID – 19 epidemic, which is causing individuals to experience a range of emotions. In the same year, [11] researchers employed machine learning models as a meta-model to examine polarity in Spanish tweets. They also looked at a variety of diseases and related-vaccines to see how stress and emotions affected users. [1] identified bots that acts as an active news promoter of the COVID – 19 tweets. The study found 127 bots which were used to extract 185,000 messages related to COVID–19 outbreak.

Another research looked at the people's sentiments about different vaccinations from companies like Pfizer, Moderna, and AstraZeneca. The paper [26] proposed a classifier utilizing the super- vised KNN algorithm based on tweets. The dataset is divided into three categories by the algorithm: positive, negative, and neutral. Thus, various models and algorithms can be used to analyze the sentiments with dataset from different vaccines.

Besides sentiments, the research has also focused on topic analysis. [12] authors focused on 530,000 original tweets in region Persian/Farsi with respect to COVID-19. The study performed topic analysis that was mainly focused on Iranians. The study analyzed and discovered 25 topics who experience home quarantine as one of the major topics. The study can be used to track public response and developments related to COVID-19. Keeping COVID-19 and health issues into consideration [13] also employed the NLP approach to perform topic analysis and used LSTM recurrent neural network to perform sentiment analysis using COVID-19 comments being posted on different social media handles.

Previous research has mostly focused on one type of approach in their work, such as detecting the sentiments of texts (an NLP-based approach) or investigating the construction of echo chambers in a network (network-based approach). However, in order to put a step towards mitigating misinformation, it is important to understand multiple aspects of online social media to get a wider picture in a single study. To be more explicit, the study of text-based features such as sentiments extracted from the texts combined with identifying the communities formed by the different types of users tells us more about the patterns of such users and, hence, combating misinformation. Furthermore, in the year 2020, when the latest vaccines were introduced by various research laboratories worldwide, people started tweeting their opinions and viewpoints over different vaccines. Therefore, analyzing the different vaccine types can provide more in-depth insights into how vaccines can affect misinformation. To summarize, in this work, we first employ a transfer learning strategy to label our curated COVID-19 vaccines dataset. Then, we apply numerous analyses (such as network and NLP-based approaches) to determine the patterns used by various types of users. Finally, we conclude our investigations along with future directions and limitations of this work.

3. Dataset

Our investigation into the debate on COVID-19 vaccines in India using social media platforms is based on the dataset collected from Twitter. The data collection is done using Twitter Streaming Application Programming Interface or Twitter API [21], to collect tweets based on specific hashtags related to various COVID-19 vaccination [4]. In this work, we focused on five well-known COVID-19 vaccines (Covaxin, Covishield, Moderna, Pfizer, and SputinkV) and we used hashtags related to these vaccines to collect the dataset. Specifically, we consider the following list of hashtags: #covaxin, #covishield, #pfizer, #modernaVaccine and #sputinkV.

In total, we gathered over 204,000 tweets, covering 3 months, from February 1st, 2021 to April 30th, 2021. This time period is important as it includes several vital real-world events in India related to COVID-19 vaccines that influenced and dictated people's discussions on Twitter. Table 1 provides an overview of our dataset, showing the total numbers of tweets per vaccine (column 2), and the total number of unique users per

Table 1

Overview of our dataset

Vaccine Name	#Tweets	#Unique Users
Covaxin	90677	40689
Covishield	69116	37236
Moderna	6387	5143
Pfizer	15082	11554
SputnikV	23492	7468

vaccine (column 3). We can observe that approximately 80% of our dataset belongs to Covaxin, and Covishield, which further indicate that the focus of this work is debate on vaccines in India.

3.1 *Temporal Evolution of Vaccines Debate*

Next, we delve deeper into the temporal evolution in the vaccines discussions by showing the daily time series of the numbers of tweets in Figure 1. Here, x-axis shows the dates from February 1st, 2021 to April 30th, 2021 and y-axis represents the number of tweets for different vaccines. The Figure 1 also highlights the days of some important real-life events related to COVID-19 vaccines. Please note that there are some significant spikes in the volume of tweets, which coincide with the identified events. On 2*nd* February 2021, Pfizer Moscow trial reported an efficacy of 91.6% after the second dose for all age groups with no unusual side effects (event #1 in the Figure 1, and Row 1 in Table 2), and on 1*st* April 2021, the third phase of vaccination of covishield and covaxin for the middle aged people i.e., age group of 45 and above was started (event #3, and Row 3 in Table 2).

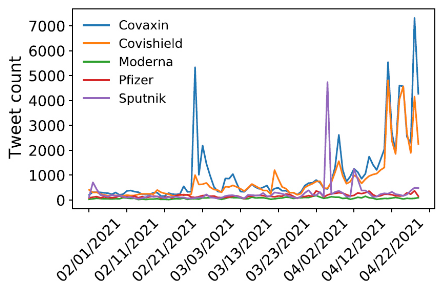


Figure 1

Temporal evolution in the vaccines related discussions.

On 5*th* April 2021, Sputnik vaccine roll-out for emergency use in India (event #4 in Figure 1, and Row 4 in Table 2). The introduction of Sputnik kicks up conversations concerning the motives, causes, and effects of Sputnik. This also sparked a controversy about (1) comparing all three vaccines in India, (2) reasons behind the hurried launch of Sputnik, and (3) scarcity of Covaxin or Covishield. On 12*th* April 2021, India approved the use of SputnikV vaccine for emergency use against COVID-19 based on strong immunogenicity evidence (event #5 in Figure 1, and Row 5 in Table 2). In addition, 62 countries had granted Sputnik V emergency usage permission. This became a significant matter of discussion, as the sputnik vaccination was now receiving certification from all over the world. Users from many nations were seen developing various groups depending on their views over the introduction of the Sputnik vaccination in their country [7, 8, 9, 10].

As the Sputnik vaccine received approvals in various countries, it became popular around the world. On 21*st* April 2021, Bharat Biotech announced phase 3 interim analysis results of Covaxin with vaccine efficacy of 78% (95%CI: 61-88) against mild, moderate, and severe COVID-19 disease (event #6 in Figure 1, and Row 6 in Table 2). The efficacy against severe COVID-19 disease was 100% (95% CI: 60-100), impacting reduction in hospitalizations. In addition, the efficacy against asymptomatic COVID-19 infection was 70%, suggesting decreased transmission in COVAXIN recipients. This bought a twist in the uprising covaxin again into picture. Users started discussing and debating about the pros and cons of different vaccines. Encouraging one another to get vaccinated with a specific vaccine while criticizing the another. On 28*th* April 2021, registration for the phase-3 began in India (event #7 in Figure 1, and Row 7 in Table 2). A single-day record of nearly 13.3 million people registered.

Table 2

Timeline events

S. No.	Date	Event
1	02/Feb	Pfizer: On 2 February 2021, an interim analysis from the Moscow trial was published in The Lancet reporting an efficacy of 91.6% (95% CI, 85.6–95.2%) after the second dose for all age groups, with no unusual side effects.
2	01/March	Phase 2 of vaccination started in India for 60+ age
3	01/April	phase -3 45+ vaccination started in India - covaxin and covishield

Contd...

4	05/April	Sputnikv rollout for emergency in India
5	12/April	India approved the use of Sputnik V vaccine for emergency use against COVID-19 based on strong immunogenicity data. As of 12 April 2021, 62 countries had granted Sputnik V emergency use authorization.
6	21/April	On April 21, 2021, Bharat Biotech announced phase 3 interim analysis results of COVAXIN. The second interim analysis results in a point estimate of vaccine efficacy of 78% (95%CI: 61-88) against mild, moderate, and severe COVID-19 disease. The efficacy against severe COVID-19 disease was 100% (95%CI: 60-100), impacting reduction in hospitalizations. In addition, the efficacy against asymptomatic COVID-19 infection was 70%, suggesting decreased transmission in COVAXIN recipients.
7	28/April	Registration for the phase 3 began on 28 April in India; a single-day record of nearly 13.3 million people registered.

3.2 *Content Overview of Vaccines Debate*

As a final overview of our dataset, we look into the overall contents of tweets related to each vaccine separately. We do so by showing in Figure 2, the word clouds created using Term Frequency–Inverse Document Frequency (TF–IDF). TF–IDF is a weighing matrix used to measure the importance of a term (count + weight) to a document in a dataset. TF-IDF comprises two metrics, named term frequency (tf) and inverse document frequency (idf). TF-IDF is represented by Equation 1.

$$tf\ idf = tf\ (t, d) \times idf\ (t, d) \tag{1}$$

Here, Term Frequency is denoted as tf and is calculated from the count (c), term (t) in document (d), and represented as tf (t, d) = ctd. TF-IDF basically assign weights to every word present in the dataset, depicting the importance of all the words individually. Thus, TF-IDF word clouds is a pictorial representation of the words of the dataset that are assigned different weights. If a wordcloud has a word with higher font size that means that the word has a comparison to others.

Text preprocessing. Here, we briefly explain the basic steps that we took in preprocessing the tweet dataset before creating TF-IDF wordclouds. We start our tweets preprocessing by removing all the URLs as they do not include any valuable information for wordclouds. Afterward, the tweets are subjected to lowercasing. The emojis, emoticon, numbers, mentions (@), and hashtags (#) are excluded due to their specific semantics in

(a) Covaxin

(b) Covishield

(c) Moderna

(d) Pfizer

(e) Sputnik

Figure 2

TF-IDF word clouds

comments and overall text. In addition, all slang text is converted into its real meaning to understand the actual context of the tweet. Furthermore, we removed every punctuation mark and a variety of different stop- words available in the nltk package. Using nltk package, we did the stemming and lemmatization of the text [43].

Figure 2 shows the word clouds for the Covaxin (Figure 2a), Covishield (Figure 2b), Moderna (Figure 2c), Pfizer (Figure 2d), and Sputnik (Figure 2e) separately. Figure 2a and Figure 2b shows that Covaxin and Covishield are talked simultaneously as both word clouds have a high weightage for each other's vaccination names the word clouds hold each other vaccines names with a high weightage. As both vaccines are developed by Indian companies or in collaboration with Indian companies, users are seen conversing about its efficacy, production, availability of doses considering both as a one. It can also be observed that the vaccines are compared by the users to justify whether a user should go for a shot of covaxin or covishield vaccine. Figure 2c depicts that Moderna vaccination has received positive feedback from users. It also shows the encouragement among users to obtain their first shot of Moderna. Similarly, the word cloud of Pfizer vaccine reflects that the users shows an inclination towards thanking Pfizer for being into existence (see Figure 2d). They also discuss the success of Pfizer, and shows enthusiasm to get their first shot of Pfizer vaccine. On one side, Pfizer vaccination is being praised alot. On the other side, few users also highlights the negative impact, allergies caused after getting Pfizer. Sputnik debate, as indicated in Figure 2e, mostly focuses on various nations where sputnik's opinions have been posted and provides updates on cases in these countries. As previously stated in Section 3.1, the sputnik vaccine was deemed an emergency vaccination for a number of nations. As a result, users were largely seen debating the efficacy of the Sputnik vaccine, its approval, cause, and worldwide implications. Thus, from the content analysis, we can conclude that while Covaxin and Covishield are compared and discussed concurrently; Moderna, Pfizer, and Sputnik existed independently. Furthermore, Moderna and Pfizer vaccines are being lauded more than others. Leaving aside the Sputnik vaccine, which shown a greater proclivity for topics concerning approvals and efficacy on a worldwide scale.

4. Dataset curation

Here, we discuss the methodology followed to curate tweets in our dataset based on their leanings towards vaccination (Section 4.1). Furthermore, in Section 4.2, we labelled users based on their tweets.

4.1 *Tweet Labelling according to user polarisation*

Our work considers user's tweet as a mean to adherence its beliefs and leaning towards a vaccine, manifesting positively as vaccination support or negatively as opposing vaccination or neutral as neither supporting nor opposing. This paper categorise tweets into three broad classes of user polarisation: "Proponent" represented by tweets in which user support vaccines (for example, “I got my first shot, feeling good! Go get it now.”), “Opponent” represented by tweets in which user oppose or doubts vaccines (for example, “Suffering from this pandemic, and now from this vaccine. Do not get it.”), and “Neutral” corresponding to the tweets in which user neither support or oppose vaccines (for example, “Go corona go”). We follow the methodology outlined in Figure 9 for tweet labelling.

Ground Truth. Initially, we follow a majority voting scheme that is explained here to manually annotate a random set of 1000 tweets into three classes (Proponent, Opponent, and Neutral). The data was annotated independently by three annotators (two bachelor students and one PhD student), each of whom were intimately familiar with the definition of all three classes. They were advised to make a decision keeping the entire text in mind, rather than the sample of the tweet. The annotator agreement between three annotators was measured by the Fleiss Kappa (0.712). This score is indicative of Good Agreement among annotators[2].

Classification Models. We considered manually annotated tweets dataset as a train and test data for three classification models:

(1) **Transformation API for XML (TRAX) Model:** This is a deep learning based language model that can be used to identify the textual context and thereby label the entire dataset of the same [3];

(2) Thus, in order to study the communities formed, the tweet categories are divided into two groups. Table 3 shows the groups formed i.e., the two categories in which the users are classified, or simply the user polarisation

Once the categories were finalized, the 1000 tweets were then considered as a training dataset for three different types of models named as:

a. TRAX Model. TRAX stands for Transformation API for XML [3]. It is another type of language model of deep learning. It is an end-to-end library which mainly focuses on the speed and clarity of the code. This model is use to identify the textual language context and thereby label the entire dataset of the same.

b. BERT Model. BERT model stands for Bidirectional Encoder Representations from transformers [16]. It is a language model which

[2] https://statistics.laerd.com/spss-tutorials/fleiss-kappa-in-spss-statistics.php

Table 3

Label Mapping for User Polarisation in 2 groups

Group-1 (First category of user separation)	Group-2 (Second category of user separation)
0: Opponent	0: Informative
1: Proponent	1: Inquisitive
2: Neutral	2: Opinionated

is designed to help computers understand any type of ambiguous language. It performs this action by surrounding the text to form a context. The framework of BERT is pre-trained using Wikipedia text. It can also be fine-tunes by appending question and answers to the dataset.

c. XLNET Model. XLNET is a natural language processing model which helps to choose in how the language modelling problem is approached [5]. It is an regressive model which generates an output of the joint probability in a sequence of tokens. This is based on the transformation architecture including reassurance.

4.2 *Results obtained*

4.2.1 Implementing TRAX Model.

Following are the results obtained from TRAX: Below Figure 3 represents the validation accuracy [14] and training loss obtained from the group – 1 results. Hence, the accuracy obtained from TRAX model is 80.1%.

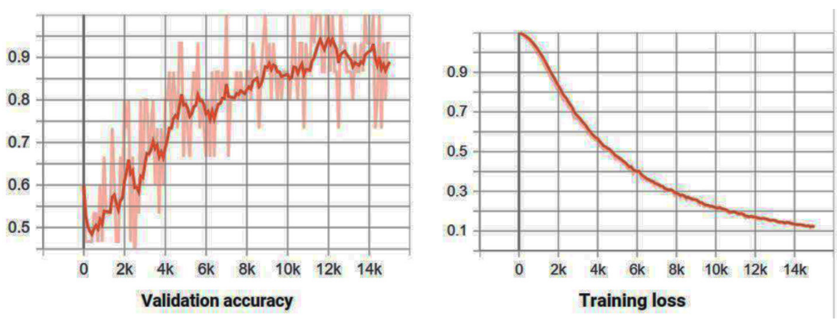


Figure 3

TRAX Model Results for Group - 1

Similarly, for group – 2, following Figure - 4 shows the accuracy and training loss curve. Which shows that the accuracy obtained by TRAX model for prediction of labels of tweets, tweeted by different users on twitter is 79.2%.

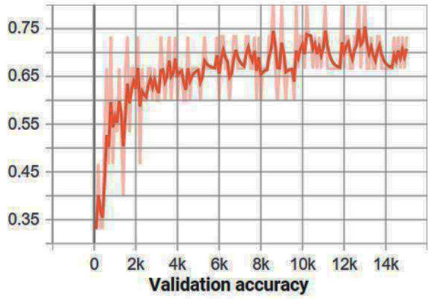
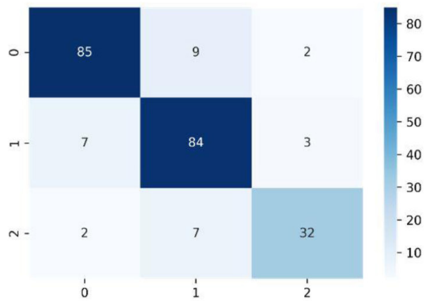


Figure 4

TRAX Model Results for Group - 2

4.2.2 Implementing BERT Model

Following are the results obtained upon testing the BERT model for labelling the dataset: Below Figure 5 represents the confusion matrix and ROC curve obtained from the group – 1 results. Hence, the accuracy obtained from BERT model is 87.0%.



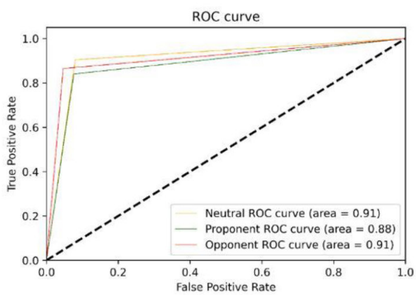


Figure 5

BERT Model Results for Group - 1

Figure 6 represents the confusion matrix and ROC curve obtained from the group – 2 results. Hence, the accuracy obtained from BERT model is 89.4%.

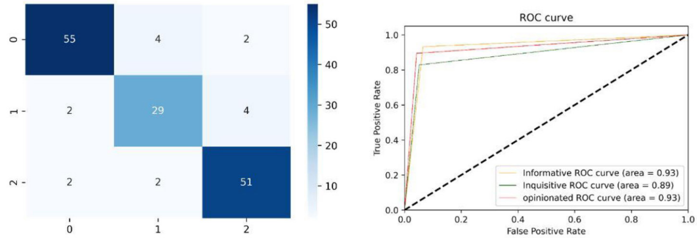


Figure 6

BERT Model Results for Group - 2

4.2.3 Implementing XLNET Model

Following are the acquired results for training and then testing the XLNET model for labelling the tweets dataset: Below Figure – 7 represents the confusion matrix and ROC curve obtained from the group – 1 results. Hence, the accuracy obtained from XLNET model is 88.2%.

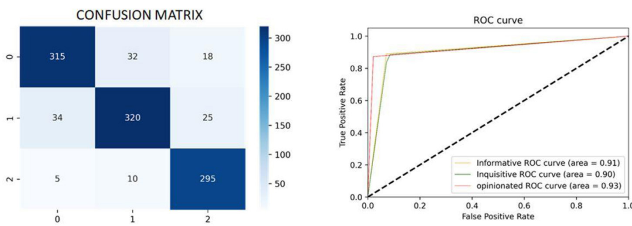


Figure 7

XLNET Model Results for Group - 1

Figure – 8 represents the confusion matrix and ROC curve obtained from the group – 2 results. Hence, the accuracy obtained from XLNET model is 90.6%.

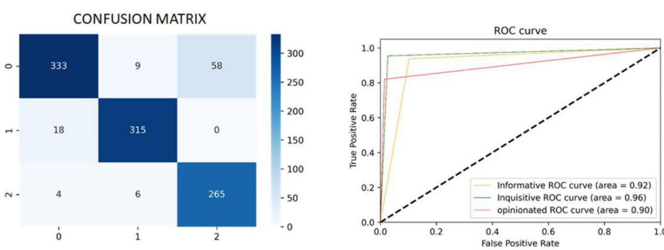


Figure 8

XLNET Model Results for Group - 2

Table 4

Comparison of Group 1 results

Model	Accuracy	Recall			Precision			F-1 Score			AUC score		
		0	1	2	0	1	2	0	1	2	0	1	2
TRAX	80.10%	0.8	0.78	0.65	0.85	0.79	0.76	0.8	0.8	0.79	0.86	0.74	0.88
BERT	87%	0.88	0.89	0.78	0.9	0.84	0.86	0.89	0.86	0.82	0.91	0.88	0.9
XLNET	88.20%	0.86	0.84	0.95	0.89	0.88	0.87	0.87	0.86	0.91	0.90	0.89	0.92

Table 5

Comparison of Group 2 results

Model	Accuracy	Recall			Precision			F-1 Score			AUC score		
		0	1	2	0	1	2	0	1	2	0	1	2
TRAX	79.2%	0.76	0.65	0.8	0.85	0.85	0.75	0.79	0.81	0.77	0.74	0.87	0.84
BERT	89.4%	0.9	0.82	0.92	0.93	0.82	0.89	0.91	0.82	0.91	0.93	0.88	0.92
XLNET	90.6%	0.83	0.94	0.96	0.93	0.95	0.82	0.88	0.95	0.88	0.91	0.96	0.90

Table 6

Final Labelled Tweet Dataset Analysis

Name of Vaccines	Unique Users	Number of Tweets					
		Informative	Inquisitive	Opinionated	Neutral	Proponent	Opponent
Covaxin	40689	36867	25059	28751	36118	48206	6353
Covishield	37236	29674	19402	20040	40702	25455	2959
Pfizer	11554	3046	2084	9952	4422	10142	518
SputnikV	7468	12393	3484	7615	17196	5319	977
Moderna	5143	1158	845	4384	1651	4450	286

Table 7

Category-wise bifurcation of Unique Users

Name of Vaccine	Number of Unique Users			
	Proponent	Opponent	Neutral	Total Users
Covaxin	17698	2047	20942	40689
Covishield	12267	1090	23878	37236
Pfizer	8090	266	3198	11554
SputnikV	1458	275	5735	7468
Moderna	3602	153	1388	5143

Thus, Table 4 and Table 5 represents the tabular representation for comparison of the results acquired by implementing three different models for labelling the entire dataset for both the classes. Thus, upon comparing XLNET model obtained the highest accuracy and therefore, was used to label the complete dataset for both the classes.

Thus, upon comparing XLNET model obtained the highest accuracy and therefore, was used to label the complete dataset for both the classes. Therefore, Table – 6 represents the final labelled dataset obtained for all the five types of vaccines dataset of tweets different users have tweeted.

5. Network Analysis

After obtaining the final dataset (ref. to section - 4, Table 6 and 7, this section focuses on the analysis withdrawn from the final dataset obtained from twitter. This section focuses on the formation of network visualizations to study the formation of different communities and groups within the twitter users based on the category of tweets they fall in.

Once the final dataset was obtained (ref. to section 3) and analyzed on different parameters. The dataset was further used to study network visualization to identify the formation of different com- munities for all the five vaccines. Network visualization refers to the graphical representation of the dataset to represent network metrics, data flows and study the interlinking of different nodes (in this case twitter users). Network visualization helps to study the formation of different groups or communities of users to identify users sharing a common thought. This also helps in distinguishing the users with different opinions forming a huge network irrespective of the place and locality they reside in. Thus, this section of the paper discusses about the network visualization for all the five vaccines along with its properties. Figure 9 shows the overall methodology used for performing network visualization.

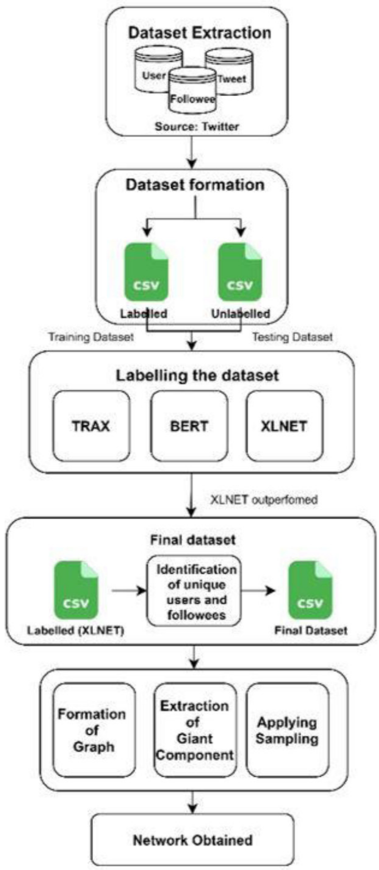


Figure 9

Methodology

Once the final dataset was obtained. Follow networks in Figure 10 and Figure 11 are constructed by extracting the followers of the unique users for all the five vaccinations. A follow network is a type of directed network which consists of users and following of the users. In this network the nodes represents the users and edge from node represents users that are following the node user. Given five different vaccine dataset, different follow networks were formed for all the five vaccines.

Following Figure 10, 11 are the two network visualizations for moderna and sputnik vaccine. In both the follow network green represents "Proponent" and red represents "Opponent". Table 8 provides the statistic of the network for all the five vaccine dataset.

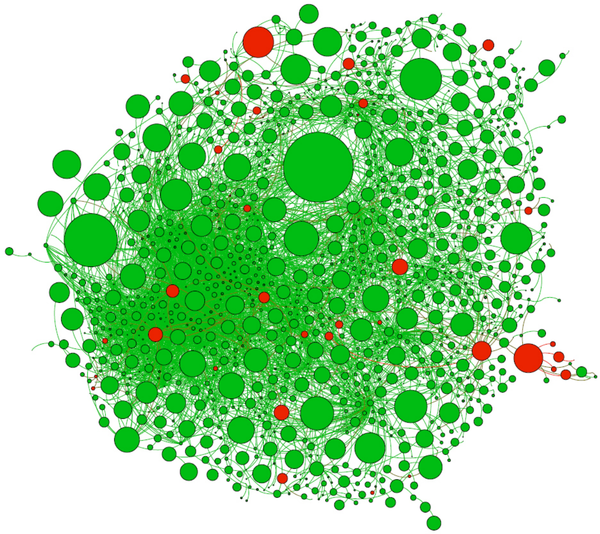


Figure 10

Follow Network - Moderna vaccine

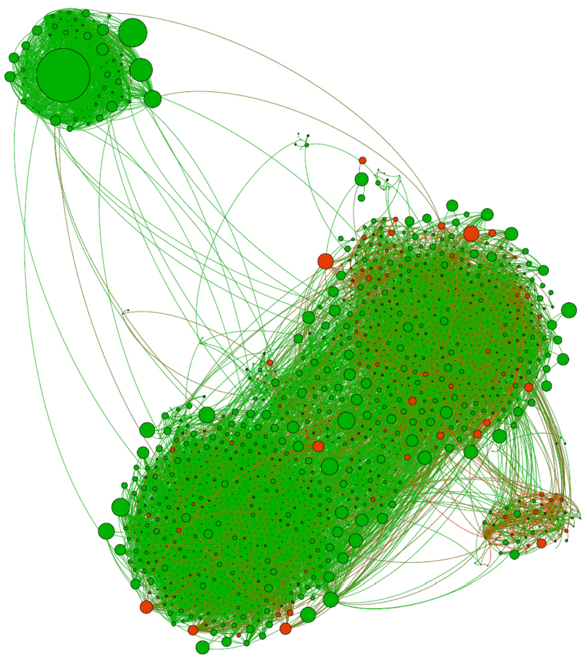


Figure 11

Follow Network - Sputnik vaccine

Table 8
Properties of Network Analysis

Properties	Covaxin	Covishield	Moderna	Pfizer	Sputnik
Number of Nodes	23648	16374	1414	4270	1789
Number of edges	876660	461871	4096	15377	11608
Average shortest path length	2.547	2.595	4.302	3.947	3.361
Diameter	9	9	12	15	10
Transitivity	0.043	0.044	0.125	0.065	0.116
Avg Clustering coefficient	0.311	0.297	0.095	0.133	0.237
Edge density	3.14E-03	3.45E-03	4.10E-03	1.69E-03	7.26E-03
Average degree	74.142	56.415	5.793	7.202	12.977
Number of Triangles	37149648	13449108	13458	98340	85629
Freeman segregation	0.180	0.166	0.475	0.173	0.373
Homophily segregation - HIa	-0.011	-0.013	0.045	0.110	0.068
Homophily segregation - HIb	0.197	0.179	0.485	0.175	0.406

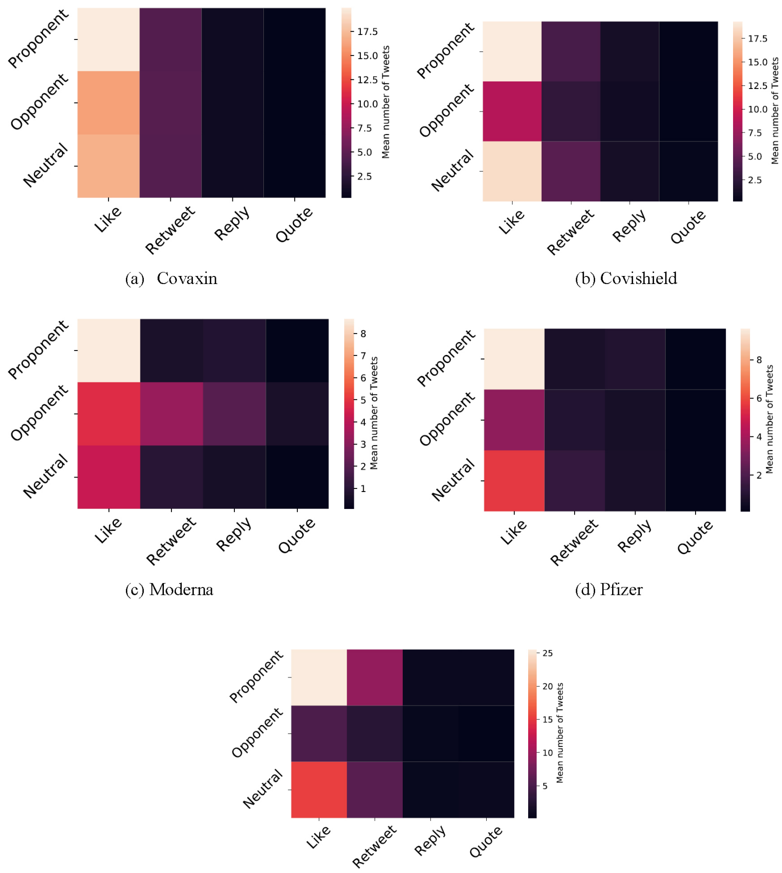
6. Results

6.1 *Engagement Analysis*

Social media users often engage with similar kinds of other users with the same belief (mindset). Specifically, certain features, such as the number of likes on a tweet and tweet replies, help each user involve and promote the same belief. Therefore, we perform engagement analysis to understand each user type's behaviour better. Figure 12 shows the engagement analysis of three types of users: proponent, opponent, and neutral on each vaccine type.

The general observations from the Figure 12 are listed below.

1. In all vaccine types, all three users prefer to use Like and Retweet over Reply and Quote.
2. While analyzing further, it can be observed that among Like and Retweet, Like is more commonly used in comparison to the other four attributes.



(a) Covaxin (b) Covishield

(c) Moderna (d) Pfizer

(e) Sputnik

Figure 12

Engagement analysis.

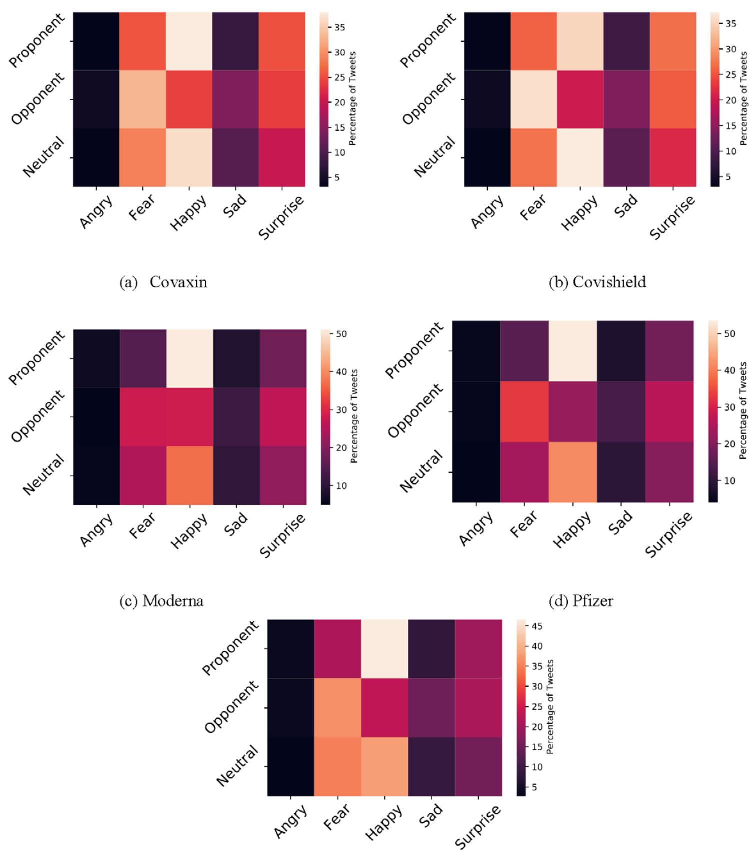
3. Regardless of vaccine, tweets concerning proponent users received the most Likes, followed by neutral (excluding Moderna vaccine) and finally opponent users' tweets.
4. Reply and Quote are similar irrespective of vaccination and tweet type (proponent, opponent, and neutral).
5. For Moderna, we observed that opponent users' tweets receive more retweets than proponent and neutral ones. In addition, the number of Likes for opponent tweets is higher than for neutral tweets.

As a result, with the exception of Moderna, we may conclude that each vaccine's engagement behavior is similar.

The next sub-section 6.2 discusses another viewpoint of analyzing the Twitter dataset obtained based on the vaccine type.

6.2 *Emotion Analysis*

Emotion analysis [2] is the process of identifying and analyzing the underlying emotions expressed in textual data. Emotion analytics can extract text data from multiple sources to analyze subjective in- formation and understand the emotions behind it [37]. Thus, in this section, we focus



(a) Covaxin (b) Covishield

(c) Moderna (d) Pfizer

e) Sputnik

Figure 13

Emotion analysis.

on the analysis withdrawn from the Twitter dataset based on the emotions of the different types of users.

Figure 13a, 13b, 13c, 13d, 13e shows the heat map of the different types of users (proponent, opponent, neutral) with varied emotions (angry, fear, happy, sad, surprise) for all the five vaccines, respectively. Following are the important findings discovered after evaluating all five vaccines heat maps for emotion analysis:

1. Angry is the least expressed emotion, followed by Sad, irre- spective of the vaccination.
2. Happy, Fear, and Surprise are the most commonly expressed emotions. Generally, Happiness triumphs over Fear and Surprise.
3. When compared to opponent users' tweets, proponent and neutral users' tweets usually exhibited higher levels of happiness (irrespective of vaccination). Similarly, when it comes to opponent, Fear is usually the most prominent emotion.
4. Users showed high Surprise for Covaxin and Covishield. (ref. to Figure 13a, 13b).
5. Users showed less Fear for Moderna and Pfizer, but high Fear for Covaxin, Covishield, and Sputnik. (ref. to Figure 13c, 13d, 13e).

6.3 *Topics Discussed around Vaccines*

Covid-19 has hit the masses mentally, physically, and financially and people are concerned not only about one but several aspects of vaccination. Being a human tendency to share your thoughts and opinion, a human being is prone to make comments and propose relative solutions to the new topics or work coming into the light. As we all know, whenever any critical event or matter comes into the limelight, we discuss it over different social media handles to know and share the views and opinions of everyone all across the world. Everyone wants to put forth their viewpoints in front of the general public to showcase their angle and views towards the topic [17]. Such discussions frequently take place on platforms like Twitter which provides citizens a stage to express their feeling, their inner voices that are very important to be heard by the entire community who is going through that critical event or situation. Therefore, in the months of February 2021, March 2021, and April 2021, when COVID-19 was at its peak, scientists and researchers from different countries came up with their vaccinations from different ends of the world. We obtained the primary topics using Latent Dirichlet Allocation (LDA) [28,29] of all the five vaccines over these three months. In addition, we calculate the coherence score [22] in order to retrieve the optimal number of topics from the dataset. Table 9 lists the topics that are discussed over three months period.

6.3.1 Covaxin Vaccine

During the rise of the pandemic situation, many researchers and scientists came up with their research for creating a vaccine to fight the COVID-19 pandemic situation. One such research was done by Indian researchers, who came up with their own vaccine to fight back coronavirus. As soon as the news came out, people started sharing their views and opinions on different social media platforms such as Twitter, Facebook, etc. Twitter was one of the platforms that held the majority of conversation over the production and introduction of covaxin.

Thus, in the month of February 2021, people started discussing over different vaccines which included covaxin too. The global discussion over different vaccines included conversations, debates, and opinions over different vaccines being produced all over the country. If we look over covaxin (ref. Table 8), in the first week of February 2021, people started discussing over the production of covaxin in India, followed by the users who started to show interest in the comparison between covaxin and covishield. Users started sharing there personal experience over the trials being held all across the country. The second week of February continued with the comparison where twitter users started the comparison of different vaccines taking health care data into consideration, followed by the needs and urgency of vaccines in different states of India. Seeing the trial of covaxin booming up, people started trusting as well as started demanding covaxin from different states. The production of covaxin started to gain with the passage of time, the twitter users showed enthusiasm towards covaxin. As covaxin doses started to roll out, many people were eager to get a shot. It was also observed that the twitter users from both the communities debated over the advantages and disadvantages over different vaccines. But at the same time the proponents of covaxin were enthusiastic for getting their first shot. By the end of the month, covaxin trials were successfully recorded and demand of covaxin rose exponentially. Twitter users from different countries showed there eagerness and enthusiasm for getting jagged with covaxin. Therefore, the supply of covaxin started across the sea.

In the beginning of the March 2021, as users were getting jagged by covaxin. They were also discussing about the side effects and benefits of covaxin over covishield. The discussion also approved the celebration of covaxin getting approvals from different countries. By the mid of march 2021, twitter users started sharing there reviews about covaxin. While covaxin was being appreciated all across the globe the production of covaxin was also taking a higher step. People started tweeting, encouraging the people to get jagged with covaxin. The supply and demand of covaxin was on the top. Whereas, on the other hand different vaccines from different countries were also being approved. By getting there approvals, in the end of week of march, users started debating over different vaccines efficiency and capabilities. Both opponents and proponents of covaxin started to debate over covaxin's efficiency with those of other vaccines. Specially covishield was being compared with covaxin due to the same production country.

The same discussion continued in the month of April 2021, people started raising questions about the effectiveness of covaxin in comparison to covishield and sputnik. By the second week of April 2021, covaxin started to gain negative publicity, which declined the patients to get a shot of covaxin. This resulted in the demand of covishield and other related vaccines. But by the end of the April month, proponents of vaccines started retweeting the importance, advantages and availability of covaxin doses, which again boosted the confidence of the people to get a shot of covaxin. The confidence in the people showed there come back for covaxin. Due to which discussion started occurring for the availability of covaxin in different countries, states of the world.

6.3.2 Covishield Vaccine

Covishield has been known as the first and foremost effective vaccine to fight against coronavirus. Since the beginning of the month of February 2021, people has been inquisitive about covishield. As soon as the production details of covishield entered into the real market users showed alot of enthusiasm about its availability of doses. Users all across the globe, started discussing and demanding covishield in their country. By the mid of February 2021, users started posting their reviews and comments about covishield which boosted the confidence of the Indian members to get their first dose of covishield. Over the next few weeks, users shared as well as discussed all the information related to covishield, starting from the availability of covishield to the nearby vaccination centers update to each

Table 9

Monthly Topic Discussion around the vaccines

Name of Vaccine	Monthly Topic Analysis			
		Topic - 1	**Topic - 2**	**Topic - 3**
Covaxin	Month - 2	Global discussion over vaccine	Trails of Covaxin in India	Production of Covaxin
	Month - 3	Efficiency of Covaxin	Availibity of Covaxin doses	Comparison between Covaxin and Covishield
	Month - 4	Availability of Covaxin doses	Production of Covaxin in different states	Efficiency of Covaxin in comparison to Covishield and Sputnik
Covishield	Month - 2	Availability of doses in other countries	Vaccination across the globe	Comparison of Covishield with different vaccines
	Month - 3	Avaliability of Covishield doses	Production and approval of different vaccines	Covishield shots taken by Indian Members
	Month - 4	Price discussion of vaccines	Comparison of Covishield with different vaccines	Availability of Covishield doses
Moderna	Month - 2	Comparison of Moderna with Pfizer	Appreciation of Moderna	Availability of doses of Moderna
	Month - 3	Comparison of Moderna with Pfizer	Availability of Moderna doses	Appericiation of Moderna
	Month - 4	Approval to Moderna vaccine	Availability of Moderna doses	Comparison of Moderna with Pfizer
Pfizer	Month - 2	Effects of Pfizer vaccine	Pfizer-related updates	Doses of Pfizer vaccine
	Month - 3	Doses of Pfizer vaccine	Effects of Pfizer vaccine	Record of shots taken of Pfizer vaccine
	Month - 4	Appreciation of Pfizer	Vaccination update of Pfizer	Doses of Pfizer vaccine

Contd...

SptunikV	Month - 2	Efficiency of Sputnik	Update on suptnik doses	Sputnik related news
	Month - 3	Global effects of covid	Sputnik as a hope	Comparison of Sputnik with other vaccines
	Month - 4	Approval of suptnik	Sputnik as a hope	Sputnik related news

other. The month of February 2021, showed the complete all the inquisitive users and proponents of covishield spread all across the market encouraging the other users to join in, in this campaign.

As February 2021, marked a new start for covishield, opponent of covishield started showing their criticism in the beginning of march. Users started comparing covishield with other vaccines such as covaxin, sputnik, etc. But the criticism did not lasted for a long time. The opponents of covishield were not able to speak up seeing the enthusiasm and confidence in among the proponents. The availability of covishield and the updates of doses were solely being discussed by the users. Thus, the entire month of march covishield proponents kept on discussing the needs and updates of covishield sharing their experience with covishield.

The demand of covishield was rising. Thus, in the month of April 2021, people started discussing over the price distribution of vaccines. Users from all across the globe started discussing covishield doses updates, avail-abilities and cost price of different vaccines comparing their demands and effectiveness in terms of the cost prize. Therefore, though the demand of covishield was rising so was the prize distribution, which started affecting the viewpoint and opinion of different users towards covishield.

6.3.3 Moderna Vaccine

Moderna is a covid-19 vaccine which has been produced after pfizer and AstraZeneca in United Kingdom. Moderna has always been the most effective vaccine. At the beginning of February 2021, modern has been discussed in the form of the symptoms people have seen after getting the shot. Moderna was one such vaccine that has been discussed about its symptoms that one person gets after getting the shot it has been also compared with Pfizer in terms of efficiency and effectiveness proponents of Moderna has also reported that it shows no side Effects while the opponents of Moderna have been criticizing it as it has a lot of side-effects in comparison to Pfizer and other vaccines. Later on, in the first week of February 2021 proponents of Moderna were eagerly waiting for the availability of doses of modern vaccine later on along with the discussion comparison of modern are vaccine with different vaccines. Results of modern vaccines showed a very good response making the trials of Moderna successful and increasing the shots of Moderna given to the people all across the globe. The symptoms and modern availability were being continuously discussed over the month of February 2021.

In the month of March 2021, users were scene appreciating moderna vaccines for such an effective response. The users started encouraging everyone to get a shot of moderna. Users also started to thank the producers of moderna. Users from different countries also started to demand moderna vaccine availability in their own country making moderna spread all across the globe. Daily up- dates of moderna vaccine was being shared by the proponents of moderna. Alot of inquisitive users were entering into the circle of getting regular updates about the availability of doses and the centers supplying the shots of moderna. Family, friends and other people started pointing the fact that moderna is the best vaccine to be rolled out and to take a shot of. By the end of the month, users started to discuss about the decline in the number of covid-19 cases after getting jagged by moderna.

The month of April 2021, users from different countries shared their update about moderna, the success of trials of moderna was being discussed with a huge enthusiasm. Users from all across the globe started appreciating the production of moderna vaccine. Some opponents of moderna also started questioning moderna's capability in comparison to pfizer and other vaccines, Various campaigns started to led the debate among the opponents and proponents of moderna vaccine. By the end of the month, users started to give some opinionated views about moderna and its effectiveness over coronavirus. The debate thus concluded by the appreciation of moderna with again the discussion over the supply and availability of doses of moderna vaccine in different regions.

6.3.4 Pfizer Vaccine

Pfizer vaccine is one of the vaccine that gained global recognition over a short duration of time. In the beginning of the February 2021, users were unaware about the efficiency of pfizer which led to the discussion over the update on pfizer doses, the availability in different countries and the symptoms of it. The effects of pfizer vaccine was discussed by the users over a long period of time, which results in the favour of pfizer.

Soon, the discussion helped pfizer gain the popularity. Users started discussing about the procedure to get vaccinated by pfizer vaccine.

Soon in the month of march 2021, pfizer started gaining proponents which resulted in the global discussion over pfizer vaccine. Users started getting shots of pfizer vaccine. Daily updates of Pfizer vaccine was being shared. Alot of inquisitive users were eager to the introduction of pfizer in their countries. Proponents of users started to share their experience by writing their reviews after getting the shot of pfizer vaccine. Pfizer was officially being appreciated all across the globe. Users started to demand more supply and production of daily doses of pfizer which resulted in the exponential growth of pfizer vaccine in different regions. By the end of march 2021, the daily updates, reviews, comments about pfizer's effectiveness started gaining popularity and users even started recommending pfizer vaccine for getting shots.

Along with the appreciation received by pfizer, some opponent users of pfizer started to lay down the comparison chart of Pfizer with those of different vaccines. Initially in the month of april 2021, opponent users started to question pfizer upon their trails and even compared Pfizer's symptoms with those of different vaccines. During the first week of april 2021, users were being misguided by the opponents of pfizer, making pfizer's demand declining day by day. But by the second week, the proponents of pfizer vaccine started their debate and campaign, encouraging the other user's to get pfizer shot. User's still appreciated the production of pfizer and thereby, making pfizer supply all across the globe.

6.3.5 Sputnik Vaccine

Sputnik was one of the vaccines manufactured by Russia's Gamaleya National Centre of Epidemiology and Microbiology. In the month of February 2021, once there was a sign of a new vaccine being rolled out by a Russian laboratory, people started expressing their views and opinions about the trials and efficacy of the vaccine via the medium of Twitter. The initial week of February, focused on the discussions all about the trials that were being conducted by the Russian laboratory for computing the results of Sputnik [24]. The week had all sorts of tweets covering the discussion of the sputnik trails, its success in the different segments. The week also collected some successful data with respect to the showing success ratio of Sputnik trails. The second week further continues the discussion by the availability of the vaccine within the country. It later on also shows some tweets where people from different countries show interest in Sputnik to be imported into their countries after seeing the success ratio of the vaccine. Moving on as soon as the availability of Sputnik was on the spot, people were found also talking about the supply and production of vaccines. They started comparing Sputnik with the existing vaccines to find the pros and cons of different vaccines. Tweets related to some direct and indirect comparisons between Sputnik and other vaccines from different countries were also seen which included factors like the origin country of the vaccine, the cost, production facilities, and many more. By the end of the month, Twitter users were observed to have gained confidence over sputnik vaccines, as there were discussions about the successful approvals of Sputnik in different countries. The tweets also reflected the enthusiasm of the twitter users for taking Sputnik as another hope to overcome this pandemic situation. From comparing Sputnik with different pre-existing vaccines users have now moved towards discussing the confidence they feel after Sputnik, its efficiency, and success ratio.

As time passed, COVID-19 was getting even stronger. The number of patients getting COVID-19 has been widely increasing which has affected the global economy to a huge extent. In the month of March, people were already suffering from the COVID-19 pandemic, wherein the issues regarding finance, economy, and mental stress also joined in. Twitter users started tweeting about the concerns and depressing state they were going through. Thus, though vaccines were being a support to the community of people, still everyone was deeply affected by the change of lifestyle, the market rates and were concerned about their families, health and even when they started questioning their existence in the next few months. In the coming week of march though, users focused on the economy, health, and global changes due to COVID-19 still there was Sputnik observed being settling in the mindset of the users. Twitter users were observed to still tweet about Sputnik as the solution to this pandemic. They started getting faith and confidence over Sputnik. Even the people outside Russia want to get the approvals of Sputnik within their own country for the betterment of the country. Users started discussing over Sputnik about its efficiency, effects, pros, cons across countries. They started gaining insights and updates on the availability of doses. Some users were also observed tweeting about the government guidelines to be updated for the sputnik to be available as soon as possible. Overall the month of march brought faith, confidence and enthusiasm among the users to fight against the pandemic by the means of sputnik.

People all across the world have started showing enthusiasm for getting sputnik valid within their own countries. Therefore, in the month of April, sputnik started getting global approval from all across the world. Twitter users started showing a great level of joy, interest and happiness towards the approval of sputnik. Seeing the successful validation of sputnik in different countries, users were observed to be more inquisitive about the production of sputnik. The trial results, efficiency, doses availability, government guidelines etc. everything has been discussed as well as updated by the users all across the country. In the mid of march, twitter users started showing their hopes from sputnik to bring a change in this pandemic situation [23]. Users started updating news, information about the latest updates on sputnik in their country. But as we know, unlike sputnik there already existed similar purpose vaccines. Twitter users started showing some group formatting for proponent and opponent categories for sputnik vaccine. It was also observed that along with fame and hope from sputnik there also existed tweets which were questioning sputnik's success and abilities to improve the current pandemic situation. Debates, difference of opinion and fights over sputnik could be clearly seen towards the end of the month of march, which concludes that every vaccine had its own proponent and opponent users, who are responsible for changing human perspective and thinking towards the vaccines[18].

7. Conclusion

Social networks have become a defacto platforms for discussing important and critical events. Often the chain of discussions lead to the creation of large corpus of data which can give new insights when large scale data analytics is performed. In other words, Social media platforms have become an important tool to analyse opinions of individuals. In this work, we studied a large scale corpus of tweets regarding various vaccines, collected over a period of 2.5 months or how many weeks? for COVID-19. Using semi-supervised process, we analysed the content (Tweets) and the content creators (Tweeter users), and explored this content in light of user polarisation. In particular, we explored NLP, social networks analysis and AI techniques for our analysis. We segment Twitter users into two categories proponents (those propelling the idea of vaccination) and opponents (those opposing the idea of vaccination). Further, we have also categorized the Twitter users on the basis of information seekers (inquisitive), information providers (informative) and opinion providers (opinionated). Our results indicated that majority of users belong to proponents and a few opponents are present in the network. Alongside, we also analyse the tweets and the results indicate that users had depicted more of happy and surprised emotion towards the vaccination. The users were giving more engagement by likes and retweets as compared to quotes and reply.

References

[1] A. Al-Rawi and V. Shukla, "Bots as active news promoters: A digital analysis of Covid-19 Tweets," *Information*, vol. 11, no. 10, p. 461 (2020).

[2] V. Gupta, S. Juyal, and Y. C. Hu, "Understanding human emotions through speech spectrograms using deep neural network," *The Journal of Supercomputing*, pp. 1-30 (2022).

[3] D. Boscolo, M. Krämer, M. Durante, M. C. Fuss, and E. Scifoni, "TRAX-CHEM: A pre-chemical and chemical stage extension of the particle track structure code TRAX in water targets," *Chemical Physics Letters*, vol. 698, pp. 11-18 (2018).

[4] V. Gupta, N. Jain, P. Katariya, A. Kumar, S. Mohan, A. Ahmadian, and M. Ferrara, "An emotion care model using multimodal textual analysis on COVID-19," *Chaos, Solitons & Fractals*, vol. 144, p. 110708 (2021).

[5] M. D. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer, "Predicting the political alignment of twitter users," in 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, pp. 192–199 (2011).

[6] A. Cossard, G. D. F. Morales, K. Kalimeri, Y. Mejova, D. Paolotti, and M. Starnini, "Falling into the echo chamber: The Italian vaccination debate on Twitter," in Proceedings of the International AAAI Conference on Web and Social Media, vol. 14, pp. 130-140 (2020).

[7] India Today Web Desk, "Phase 3 vaccination starts today to cover all aged 45+ as India's Covid figures soar again | 10 points," India Today, [Online]. Available: https://www.indiatoday.in/coronavirus-outbreak/story/phase-3-covid19-vaccination-starts-today-india-coronavirus-case-situation-1785780-2021-04-01.

[8] The Hindu Net Desk, "Morning Digest | Nod soon for Sputnik V rollout, amid virus surge; Night curfew, weekend lockdown announced

in Maharashtra, and more," The Hindu, [Online]. Available: https://www.thehindu.com. [Accessed: Dec. 19, 2024].

[9] The Hindu Net Desk, "Phase 2 of vaccination begins for senior citizens and those above 45 years with co-morbidities," The Hindu, [Online]. Available: https://www.thehindu.com/news/national/coronavirus-live-updates-march-1-2021/article33958676.ece. [Accessed: Dec. 19, 2024].

[10] D. Ghosh, "Around 1.33 Crore Apply For Covid Vaccination: 10 Points," NDTV, [Online]. Available: https://www.ndtv.com/india-news/vaccinatoion-cowin-27-lakh-hits-per-minute-government-sources-on-vaccine-registration-2423760. [Accessed: Dec. 19, 2024].

[11] A. Rodríguez González, J. M. Tuñas, D. Fernandez Peces-Barba, E. Menasalvas Ruiz, A. Jaramillo, M. Cotarelo, A. Conejo, A. Arce, and A. Gil, "Creating a metamodel based on machine learning to identify the sentiment of vaccine and disease-related messages in Twitter: The MAVIS study," in 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), pp. 245–250 (2020).

[12] P. Hosseini, P. Hosseini, and D. A. Broniatowski, "Content analysis of Persian/Farsi Tweets during COVID-19 pandemic in Iran using NLP," arXiv preprint arXiv:2005.08400 (2020).

[13] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: NLP using LSTM recurrent neural network approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2733–2742 (2020).

[14] P. Kellman and M. S. Hansen, "T1-mapping in the heart: accuracy and precision," *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 1, pp. 1–20 (2014).

[15] L. Krempel, "Network visualization," in The SAGE Handbook of Social Network Analysis, pp. 558–577 (2011).

[16] A. H. F. Laender, B. A. Ribeiro-Neto, A. S. Da Silva, and J. S. Teixeira, "A brief survey of web data extraction tools," *ACM Sigmod Record*, vol. 31, no. 2, pp. 84–93 (2002).

[17] H. Li and K. Yamanishi, "Topic analysis using a finite mixture model," *Information Processing & Management*, vol. 39, no. 4, pp. 521–541 (2003).

[18] D. Y. Logunov, I. V. Dolzhikova, D. V. Shcheblyakov, A. I. Tukhvatulin, O. V. Zubkova, A. S. Dzharullaeva, A. V. Kovyrshina, N. L. Lubenets, D. M. Grousova, A. S. Erokhova, et al., "Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: An interim analysis of a randomised controlled phase 3 trial in Russia," The Lancet, vol. 397, no. 10275, pp. 671–681 (2021).

[19] M. O. Lwin, J. Lu, A. Sheldenkar, P. J. Schulz, W. Shin, R. Gupta, and Y. Yang, "Global sentiments surrounding the COVID-19 pandemic on Twitter: Analysis of Twitter trends," *JMIR Public Health and Surveillance*, vol. 6, no. 2, e19447 (2020).

[20] J. Moody, D. McFarland, and S. Bender-deMoll, "Dynamic network visualization," *American Journal of Sociology*, vol. 110, no. 4, pp. 1206–1241 (2005).

[21] D. Murthy, Twitter. Cambridge, UK: Polity Press (2018).

[22] D. Newman, E. V. Bonilla, and W. Buntine, "Improving topic coherence with regularized topic models," in Advances in Neural Information Processing Systems, vol. 24, pp. 496–504 (2011).

[23] P. Raghavan, "Panel clears Sputnik, could soon be India's 3rd shot against Covid-19 surge," The Indian Express, [Online]. Available: https://indianexpress.com/article/india/sputnik-v-covid-vaccine-india-7270310/. [Accessed: Dec. 19, 2024].

[24] Reuters and K. N. Das, "India approves Russia's Sputnik V COVID-19 vaccine," Reuters, [Online]. Available: https://www.reuters.com/world/india/india-approves-russias-sputnik-v-vaccine-economic-times-2021-04-12/. [Accessed: Dec. 19, 2024].

[25] J. C. M. Serrano, O. Papakyriakopoulos, and S. Hegelich, "NLP-based feature extraction for the detection of COVID-19 misinformation videos on YouTube," in Proceedings of the 1st Workshop on NLP for COVID-19 at ACL (2020).

[26] M. F. M. J. Shamrat, S. Chakraborty, M. M. Imran, J. N. Muna, M. M. Billah, P. Das, and O. M. Rahman, "Sentiment analysis on Twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1, pp. 463–470 (2021).

[27] V. Gupta, P. Dass, and R. Arora, "Pendulating or resonating? A case of echo-chambers in Twitter," *Journal of Discrete Mathematical Sciences and Cryptography,* vol. 25, no. 1, pp. 231–240 (2022).

[28] P. Stenetorp, S. Pyysalo, G. Topić, T. Ohta, S. Ananiadou, and J. Tsujii, “BRAT: A web-based tool for NLP-assisted text annotation,” in Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, pp. 102–107 (2012).

[29] X. Wei and W. B. Croft, “LDA-based document models for ad-hoc retrieval,” in Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 178–185 (2006).

[30] J. Xue, J. Chen, C. Chen, C. Zheng, S. Li, and T. Zhu, “Public discourse and sentiment during the COVID-19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter,” PloS One, vol. 15, no. 9, e0239441 (2020).

[31] J. S. Zhang, B. C. Keegan, Q. Lv, and C. Tan, “A tale of two communities: Characterizing Reddit response to COVID-19 through /r/China_Flu and /r/Coronavirus,” (2020).

[32] V. Singh, R. C. Poonia, S. Kumar, P. Dass, P. Agarwal, V. Bhatnagar, and L. Raja, “Prediction of COVID-19 coronavirus pandemic based on time series data using Support Vector Machine,” *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 8, pp. 1583–1597 (2020).

[33] V. Bhatnagar, R. C. Poonia, P. Nagar, S. Kumar, V. Singh, L. Raja, and P. Dass, “Descriptive analysis of COVID-19 patients in the context of India,” *Journal of Interdisciplinary Mathematics*, vol. 24, no. 3, pp. 489–504 (2021).

[34] V. K. Verma, M. Pandey, T. Jain, and P. K. Tiwari, “Dissecting word embeddings and language models in natural language processing,” *Journal of Discrete Mathematical Sciences and Cryptography,* vol. 24, no. 5, pp. 1509–1515 (2021).

[35] G. S. Sharma, M. Tomar, J. Kumari, and P. Kumar, “Social media as a social capital for academic research,” *Journal of Information and Optimization Sciences*, vol. 43, no. 6, pp. 1487–1498 (2022).

[36] S. Dhingra, R. Arora, P. Katariya, A. Kumar, V. Gupta, and N. Jain, “Understanding emotional health sustainability amidst COVID-19 imposed lockdown,” in Sustainability Measures for COVID-19 Pandemic, pp. 211–235 (2021).

[37] C. Qian, N. Mathur, N. H. Zakaria, R. Arora, V. Gupta, and M. Ali, "Understanding public opinions on social media for financial sentiment analysis using AI-based techniques," *Information Processing & Management*, vol. 59, no. 6, 103098 (2022).

[38] S. Sharma, R. Sharma, and A. Datta, "(Mis)leading the COVID-19 vaccination discourse on Twitter: An exploratory study of infodemic around the pandemic," *IEEE Transactions on Computational Social Systems* (2022).

[39] R. Goel and R. Sharma, "Studying leaders & their concerns using online social media during the times of crisis-A COVID case study," *Social Network Analysis and Mining*, vol. 11, pp. 1–12 (2021).

[40] A. Chaudhary, V. Gupta, N. Jain, and K. C. Santosh, "COVID-19 on air quality index (AQI): A necessary evil?," in COVID-19: Prediction, Decision-Making, and its Impacts, pp. 127–137 (2021).

[41] V. Gupta, N. Jain, D. Virmani, S. Mohan, A. Ahmadian, and M. Ferrara, "Air and water health: Industrial footprints of COVID-19 imposed lockdown," *Arabian Journal of Geosciences*, vol. 15, no. 8, 687 (2022).

[42] V. Gupta, K. C. Santosh, R. Arora, T. Ciano, K. S. Kalid, and S. Mohan, "Socioeconomic impact due to COVID-19: An empirical assessment," *Information Processing & Management*, vol. 59, no. 2, 102810 (2022).

[43] V. Gupta, V. K. Singh, P. Mukhija, and U. Ghose, "Aspect-based sentiment analysis of mobile reviews," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 4721–4730 (2019).

[44] V. Bansal, M. Tyagi, R. Sharma, V. Gupta, and Q. Xin, "A Transformer-based approach for abuse detection in code-mixed Indic languages," *ACM Transactions on Asian and Low-Resource Language Information Processing* (2022).

Received May, 2023